

**EVALUASI DISTRIBUSI HASIL SELEKSI  
PPDB SMA NEGERI DI KOTA SURAKARTA  
MENGUNAKAN METODE DECISION TREE**



**SKRIPSI**

Disusun sebagai salah satu syarat menyelesaikan Program Studi Strata I  
pada Program Studi Teknik Informatika Fakultas Komunikasi dan Informatika  
Universitas Muhammadiyah Surakarta

**Oleh :**

**AZIZ RIDHA UTAMA**

**NIM : L 200 100 135**

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS KOMUNIKASI DAN INFORMATIKA  
UNIVERSITAS MUHAMMADIYAH SURAKARTA**

**2014**

## **HALAMAN PERSETUJUAN**

Skripsi dengan judul

**EVALUASI DISTRIBUSI HASIL SELEKSI  
PPDB SMA NEGERI DI KOTA SURAKARTA  
MENGUNAKAN METODE *DECISION TREE***

Ini telah diperiksa dan disetujui dalam sidang pendadaran

Hari : Sabtu

Tanggal : 5 Juli 2014

Pembimbing



Nurgiyatna, M.Sc, Ph.D

NIK : 881

## LEMBAR PENGESAHAN

### EVALUASI DISTRIBUSI HASIL SELEKSI PPDB SMA NEGERI DI KOTA SURAKARTA MENGUNAKAN METODE *DECISION TREE*

Dipersiapkan dan disusun oleh

**AZIZ RIDHA UTAMA**

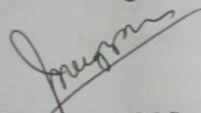
NIM : L 200 100 135

Telah dipertahankan di depan Dewan Penguji

pada tanggal 5 Juli 2014

#### Susunan Dewan Penguji

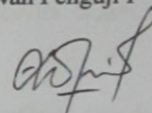
Pembimbing I



Nurgiyatna, M.Sc, Ph.D

NIK : 881

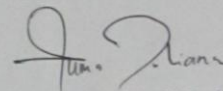
Dewan Penguji I



Aris Rakhmadi, S.T., M.Eng.

NIK : 983

Dewan Penguji II



Irma Yuliana, S.T., M.M.

NIK :

Skripsi ini telah diterima sebagai salah satu persyaratan  
untuk memperoleh gelar sarjana

Dekan

Fakultas Komunikasi dan Informatika



Husni Thamrin, S.T, M.T, Ph.D

NIK : 706

Ketua Program Studi

Teknik Informatika



Dr. Heru Supriyono, M.Sc

NIK : 970

## DAFTAR KONTRIBUSI

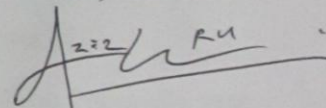
Dengan ini saya menyatakan bahwa skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Berikut saya sampaikan daftar kontribusi dalam penyusunan skripsi :

1. Nurgiyatna, M.Sc, Ph.D, selaku dosen pembimbing yang memberikan masukan dan arahan penyusunan skripsi ini.
2. Saya sendiri yang mengolah data dan penghitungan metode *decision tree* berdasarkan sumber-sumber pada buku, literatur dan *internet*.
3. Program aplikasi yang saya gunakan dalam pembuatan skripsi ini adalah Sistem Operasi Windows 7, Microsoft Word, Microsoft Excel, Rapid Miner.
4. Saya menggunakan laptop dengan spesifikasi Prosesor Intel®Core™ i5-2410M, 2.30GHz, RAM 4GB, dan Hardisk 640GB dalam pembuatan skripsi ini.

Dengan demikian pernyataan dan daftar kontribusi saya buat dengan sejujurnya. Saya bertanggung jawab atas isi dan kebenaran daftar diatas.

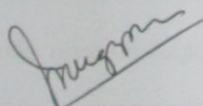
Surakarta, Juli 2014



Aziz Ridha Utama

Mengetahui:

Pembimbing I



Nurgiyatna, M.Sc, Ph.D

NIK : 881

## **MOTTO DAN PERSEMBAHAN**

### **MOTTO :**

“Kerjakanlah pekerjaanmu dengan niat tulus dan penuh keiklasan,

Maka akan kamu terima hasil yang memuaskan,

Jika kamu mengerjakan dengan keterpaksaan

Maka hasilnya pun akan berantakan “

(Aziz Ridha Utama)

### **PERSEMBAHAN :**

1. Kedua orang tuaku tercinta, Bapak Lilik Suparwanto dan Ibu Siti Umami untuk kasih sayang, doa dan dukungan moral hingga materi serta memberikan kesempatan saya untuk menempuh pendidikan sampai saat ini.
2. Adik - adikku Lisya Nurlaily Hajar Marfuah dan Fadhilah Pralampita yang selalu mendukung dan menjadi semangatku untuk menjadi lebih baik.
3. Teman - teman GEMBEL ( Gerombolan Belajar ), Yakka, Jati, Fikri, Adjie, Lukman, Amin, Eko, Bias, Rizal, Agung, dan Alvin teman seperjuangan baik suka maupun duka selama 4 tahun ini.
4. Teman - teman kelas E Informatika angkatan 2010, teman menuntut ilmu dari awal masa perkuliahan.

5. Teman - teman HIMATIF UMS, tempat pertama kali saya mengenal organisasi.
6. Semua pihak yang selalu ada untuk mendukung saya yang tidak bisa kusebutkan satu persatu.

## KATA PENGANTAR

Alhamdulillah, kami panjatkan puji syukur kehadiran Allah SWT yang telah melimpahkan rahmat, nikmat, dan hidayah-Nya sehingga penulis dapat menyelesaikan skripsi ini dengan judul “ Evaluasi Distribusi Hasil Seleksi PPDB SMA Negeri Di Kota Surakarta Menggunakan Metode *Decision Tree* ”.

Skripsi ini disusun untuk memenuhi kurikulum pada Program Studi Teknik Informatika Universitas Muhammadiyah Surakarta, sebagai kewajiban mahasiswa dalam rangka menyelesaikan program sarjana.

Penulis menyadari bahwa skripsi ini jauh dari kesempurnaan, oleh karena itu kritik saran yang membangun dari berbagai pihak sangat diharapkan penulis demi perbaikan - perbaikan ke depan.

Terwujudnya skripsi ini tidak lepas dari bantuan berbagai pihak. Oleh karena itu dalam kesempatan ini penulis ingin mengucapkan terima kasih yang sebesar - besarnya kepada :

1. Allah SWT yang telah memberikan nikmat dan hidayahnya kepada penyusun sehingga skripsi ini dapat selesai.
2. Kepada orang tua yang selalu memberikan dukungan berupa doa, semangat, dan motivasi dengan tiada hentinya kepada penulis.
3. Bapak Husni Thamrin, S.T, M.T., Ph.D. selaku Dekan Fakultas Komunikasi dan Informatika Universitas Muhammadiyah Surakarta.
4. Bapak Dr. Heru Supriyono, S.T., M.Sc selaku Ketua Jurusan Teknik Informatika Universitas Muhammadiyah Surakarta.

5. Bapak Nurgiyatna, M.Sc, Ph.D selaku pembimbing skripsi yang telah memberikan bimbingan, dukungan, dan pengarahan kepada penulis sehingga dapat menyelesaikan penyusunan skripsi dari awal hingga akhir.
6. Kepala Kantor Kesatuan Bangsa dan Politik Kota Surakarta yang memberikan kemudahan dalam proses perizinan penelitian.
7. Kepala Badan Perencanaan dan Pembangunan Daerah Kota Surakarta yang memberikan kemudahan dalam proses perizinan penelitian.
8. Kepala Dinas Pendidikan Pemuda dan Olahraga Kota Surakarta yang memberikan perizinan tempat untuk melakukan penelitian.
9. Segenap dosen dan karyawan program studi Teknik Informatika atas bantuan dan ilmu yang diberikan kepada penulis selama masa perkuliahan dari pertama masuk kuliah hingga dinyatakan mendapat gelar Strata 1.
10. Semua pihak yang tidak bisa disebutkan satu - persatu yang telah membantu dari awal hingga terselesaikannya skripsi ini.

Akhirnya penulis berharap semoga skripsi ini dapat berguna untuk semua pihak dan bermanfaat bagi penulis, serta pembaca pada umumnya dalam menambah pengetahuan dan wawasan ilmu. Aamiin.

Surakarta, Juli 2014

Aziz Ridha Utama



## DAFTAR ISI

<b>Halaman Judul .....</b>	<b>i</b>
<b>Halaman Persetujuan .....</b>	<b>ii</b>
<b>Halaman Pengesahan .....</b>	<b>iii</b>
<b>Daftar Kontribusi .....</b>	<b>iv</b>
<b>Motto dan Persembahan .....</b>	<b>v</b>
<b>Kata Pengantar .....</b>	<b>vii</b>
<b>Daftar Isi .....</b>	<b>ix</b>
<b>Daftar Tabel .....</b>	<b>xii</b>
<b>Daftar Persamaan .....</b>	<b>xiii</b>
<b>Daftar Lampiran .....</b>	<b>xiv</b>
<b>Daftar Gambar .....</b>	<b>xv</b>
<b>Abstraksi .....</b>	<b>xvi</b>
<b>BAB I     PENDAHULUAN .....</b>	<b>1</b>
A. Latar Belakang .....	1
B. Rumusan Masalah .....	3
C. Batasan Masalah .....	4
D. Tujuan Penelitian .....	5
E. Manfaat Penelitian .....	6
F. Sistematika Penulisan .....	6
<b>BAB II     TINJAUAN PUSTAKA .....</b>	<b>8</b>
A. Telaah Penelitian .....	8

B. Landasan Teori .....	12
1. <i>Data Mining</i> .....	12
2. <i>Proses Data Mining</i> .....	14
3. <i>Metode Klasifikasi</i> .....	17
4. <i>Decision Tree</i> .....	18
5. <i>Algoritma ID3</i> .....	19
6. <i>Atribut Dominan</i> .....	21
7. <i>Rapidminer 5</i> .....	22
<b>BAB III   METODOLOGI PENELITIAN .....</b>	<b>23</b>
A. Waktu dan Tempat .....	23
B. Peralatan .....	23
C. Metode Penelitian .....	24
D. Diagram Alur Penelitian .....	25
E. Langkah Penelitian .....	28
1. Identifikasi Masalah .....	28
2. Penentuan Atribut Dominan .....	29
3. Penentuan Sampel .....	30
4. Penentuan Metode <i>Decision Tree</i> .....	30
<b>BAB IV   HASIL DAN PEMBAHASAN .....</b>	<b>32</b>
A. Hasil Penelitian .....	32
B. Analisa dan Pembahasan .....	32
1. Analisa Manual <i>Decision Tree</i> Menggunakan	
Algoritma ID3 .....	32

a. Penentuan Atribut .....	32
b. Perkiraan Penentuan Jumlah Sampel Data .....	33
c. Pengolahan Data .....	34
d. Implementasi dan Perhitungan Manual <i>Decision Tree</i>	
<i>Algoritma ID3</i> .....	37
1. Menentukan <i>Root Node</i> .....	38
2. Menentukan <i>Internal Node</i> .....	43
3. Menentukan <i>Leaf Node</i> .....	63
2. Implementasi <i>Decision Tree Algoritma ID3</i> Menggunakan	
<i>Rapid Miner</i> .....	65
<b>BAB V    PENUTUP .....</b>	<b>87</b>
A. Kesimpulan .....	87
B. Saran .....	89
<b>Daftar Pustaka .....</b>	<b>90</b>
<b>Lampiran .....</b>	

## DAFTAR TABEL

<b>Tabel 4.1.</b> Pengkategorian Objek .....	36
<b>Tabel 4.2.</b> Pemodelan Tabel Aturan Dari Decision Tree . ....	78

## DAFTAR PERSAMAAN

<b>Persamaan (2.1)</b> Rumus <i>Entrophy</i> .....	20
<b>Persamaan (2.2)</b> Rumus <i>Information Gain</i> .....	20
<b>Persamaan (3.1)</b> Rumus Penentuan Sampel .....	30

## DAFTAR LAMPIRAN

<b>Lampiran 1.</b> Surat Izin Penelitian DISDIKPORA .....	93
<b>Lampiran 2.</b> <i>Sampel</i> Data Awal . .....	94
<b>Lampiran 3.</b> Data <i>Training</i> .....	104
<b>Lampiran 4.</b> Tampilan <i>Graph View Decision Tree</i> Hasil Seleksi PPDB Pada Aplikasi Rapid Miner .....	114

## DAFTAR GAMBAR

<b>Gambar 2.1.</b> Tahap-Tahap <i>Data Mining</i> . .....	14
<b>Gambar 3.1.</b> Alur penelitian .....	26
<b>Gambar 4.1.</b> Sebagian Sampel Data Awal .....	35
<b>Gambar 4.2.</b> Sampel Data Setelah Pengkategorian .....	37
<b>Gambar 4.3.</b> Root Node .....	43
<b>Gambar 4.4.</b> Internal Node Tahap Pertama .....	54
<b>Gambar 4.5.</b> Internal Node Tahap ke - Dua .....	59
<b>Gambar 4.6.</b> Internal Node Tahap ke - Tiga.....	63
<b>Gambar 4.7.</b> <i>Leaf Node</i> Pada Salah Satu <i>Internal Node</i> .....	65
<b>Gambar 4.8.</b> Penyediaan <i>New Repository</i> .....	66
<b>Gambar 4.9.</b> Input <i>Data Training</i> Menuju <i>Repository</i> .....	67
<b>Gambar 4.10.</b> <i>Data View</i> Dari <i>Data Training</i> .....	68
<b>Gambar 4.11.</b> <i>Meta Data View</i> Dari <i>Data Training</i> .....	68
<b>Gambar 4.12.</b> <i>Plot View</i> Dari <i>Data Training</i> Model 1 .....	69
<b>Gambar 4.13.</b> <i>Plot View</i> Dari <i>Data Training</i> Model 2 .....	70
<b>Gambar 4.14.</b> <i>Plot View</i> Dari <i>Data Training</i> Model 3 .....	71
<b>Gambar 4.15.</b> Proses Menghubungkan <i>Port</i> Pada Setiap Model .....	73
<b>Gambar 4.16.</b> <i>Graph View Decision Tree</i> .....	73
<b>Gambar 4.17.</b> <i>Text view decision tree</i> .....	74

## Abstraksi

Pesatnya perkembangan teknologi informasi mampu membuat suatu sistem informasi dapat menyimpan berbagai jenis data dalam jumlah besar. Apabila data - data tersebut bersifat penting maka data tersebut dapat dimanfaatkan untuk membentuk suatu informasi yang memiliki nilai guna. Salah satu contohnya adalah banyaknya data hasil seleksi Penerimaan Peserta Didik Baru (PPDB) tingkat SMA Negeri di Kota Surakarta yang dimiliki oleh Dinas Pendidikan Pemuda dan Olahraga Kota Surakarta. Cara mengatasi permasalahan pembetulan informasi dari kumpulan data tersebut adalah dengan cara menerapkan proses *data mining*. Proses data mining tersebut dengan menggunakan metode *decision tree*.

Adapun metode yang dipakai dalam penelitian adalah decision tree dengan langkah - langkah sebagai berikut yang pertama adalah studi berbagai literatur yang berhubungan dengan proses pencarian informasi dari suatu kumpulan data (*data mining*). Kemudian pemilihan obyek penelitian yaitu implementasi decision tree pada hasil seleksi PPDB di Kota Surakarta dengan pertimbangan jumlah data yang ada banyak serta belum ada yang melakukan penelitian dengan menggunakan topik yang sama. Tahapan selanjutnya adalah pengumpulan data dari Dinas Pendidikan Pemuda dan Olahraga Kota Surakarta untuk dijadikan *data training*. Kemudian penentuan atribut yang diperlukan untuk pembentukan *decision tree* dari data tersebut. Tahap terakhir adalah melakukan analisa data yaitu baik dengan bantuan *software Rapid Miner 5* serta melakukan analisa perhitungan secara manual dengan metode Decision Tree menggunakan algoritma ID3 dengan mencari nilai *entropy* serta *information gain*.

Berdasarkan training dan pengujian kemudian dilakukan analisis maka hasil pembentukan decision tree menunjukkan bahwa adanya suatu atribut yang mempunyai dominasi atau pengaruh cukup tinggi terhadap hasil seleksi PPDB, atribut tersebut adalah akumulasi nilai mata pelajaran. Kemudian ketepatan dalam pembuatan *decision tree* juga berdasarkan seberapa banyak *data training* yang digunakan.

Kata kunci : *Data Mining, Decision Tree, ID3, Data Training, Entropy, Information Gain*